Volume 316, number 1, 53–58 FEBS 12011 January 1993
© 1993 Federation of European Biochemical Societies 00145793/93/$6.00

The C-terminal tripeptide of glycosomal phosphoglycerate kinase is both necessary and sufficient for import into the glycosomes of *Trypanosoma brucei*

Jürg M. Sommer, Gregory Peterson*, Gilbert-A. Keller**, Marilyn Parsons* and C.C. Wang

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, USA

Received 23 November 1992

Glycosomal phosphoglycerate kinase (gPGK) of *Trypanosoma brucei* differs from the cytoplasmic isozyme (cPGK) in its higher isoelectric point characterized by clusters of positive charges along the polypeptide chain, and a 20 amino acid C-terminal extension ending in serine-serine-leucine (SSL). While a C-terminal SSL tripeptide is apparently not capable of directing luciferase to the peroxisomes in mammalian cells [J. Cell Biol. 108 (1989), 1657–1664], we show here that it is sufficient for the import of luciferase as well as an unrelated protein, β-glucuronidase, into the glycosomes of *T. brucei*, as determined by immunoelectron microscopy. The analysis of luciferase–gPGK fusion proteins indicates that the only targeting signal for import of gPGK into the glycosome resides in this C-terminal SSL sequence.

Glycosome biogenesis; Glycosomal targeting signal; Protein sorting; Protein import; *Trypanosoma brucei*

1. INTRODUCTION

Glycosomes are a form of microbodies found among all kinetoplastids, including *Trypanosoma brucei*. Like peroxisomes, they compartmentalize specific enzyme pathways, but they are unique in *T. brucei* in that they contain most of the glycolytic enzymes commonly found in the cytosol of other living organisms. The high concentration of these enzymes in the glycosome increases the rate of glycolysis to the point where the bloodstream form of *T. brucei* can rely solely on glycolysis for its energy supply [1,2]. Many of the characteristics of glycosomes and peroxisomes have been apparently conserved during evolution [5,6]. Thus, glycosomal proteins are encoded by nuclear genes and imported post-translationally without any apparent proteolytic modification [7]. A peroxisomal protein import mechanism that depends on a C-terminal SKL-like tripeptide [8] may also function in the import of proteins into the glycosomes of *T. brucei* [9,10]. However, in a recent detailed study of the sequence requirements for in vivo import of a peroxisomal protein, firefly luciferase [11], into the glycosomes of *T. brucei* we have shown that the C-terminal targeting signal that directs proteins to the glycosome may be far more degenerate than the signal required for peroxisomal protein import [10].

T. brucei glycosomal phosphoglycerate kinase (gPGK) and the cytoplasmic isozyme (cPGK) share 93% amino acid identity [12]. The major differences between the two proteins are their isoelectric points (9.3 and 6.8 for gPGK and cPGK, respectively), a 20 amino acid extension ending in serine-serine-leucine (SSL) at the C-terminus of gPGK, as well as clusters of positive charges postulated to be on the surface of the gPGK 3-dimensional structure [13]. Although the positively charged surface clusters (hotspots) were originally proposed to contain glycosomal targeting signals, later results suggested that hotspots are not unique to glycosomal enzymes [14], nor are they present on all glycosomal enzymes. Preliminary data suggesting that a glycosomal targeting signal may be located near the C-terminus of gPGK was obtained by Fung and Clayton [9], when they fused the last 21 amino acids of gPGK to the C-terminus of chloramphenicol acetyltransferase (CAT) and observed in a transient *T. brucei* transformation system that a substantial portion of the CAT activity cosedimented with the glycosomal fraction on a sucrose density gradient. However, an exact localization of the targeting signal, as well as direct proof of glycosomal import of such fusion protein, have not been available until the recent findings that a C-terminal SSL tripeptide was among the sequences that can target luciferaseΔSKL to the glycosome in stable transformants of *T. brucei* [10]. Although the same chimeric protein was shown previously not to be im-

Correspondence address: J. Sommer, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143-0446, USA. Fax: (1) (415) 476 3382.

**Present address:* Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109, USA.

***Present address:* Department of Preclinical Pharmacology, Genentech Inc., South San Francisco, California 94080, USA.

ported into the mammalian peroxisome [8], these findings led us to propose that the SSL tripeptide sequence may constitute the entire glycosomal targeting signal of gPGK [10]. However, our studies did not rule out the possible involvement of other position(s) in the polypeptide chain, which may be shared by gPGK and luciferase, in the glycosomal import. In this study, we have analyzed fusion proteins of either luciferase⊿SKL or β-glucuronidase with gPGK and cPGK sequences to show that the C-terminal SSL sequence is both essential and sufficient for targeting gPGK to the glycosomes of *T. brucei*.

2. MATERIALS AND METHODS

To produce the transformation vector pLUH206 (Fig. 1), the luciferase gene was inserted into pBluescript (Stratagene, San Diego, CA) for production of single-stranded DNA (+ strand) and annealed with the mutagenesis primer 5′CTGAATACAGTTACTTACTTACG-TAGGCCTGCGGCC-GCCCTTCTTG. Mutated plasmids were recovered as described [10] and the luciferase gene was inserted into the multicloning site of pTSA-HYG2 on a *Xho*I to *Sma*I fragment. The vector pTSA-HYG2 was made by deleting a 870 bp *Nru*I to *Hpa*I fragment containing non-essential SV40 sequences downstream of the tubulin intergenic region from the vector pTSA-HYG1 [10]. The luciferase mutant ending in SSL has been described previously [10]. The *T. brucei* gPGK and cPGK genes and the 60 bp C-terminal extension of gPGK were flanked with *Not*I sites by the polymerase chain reaction (PCR) method. The PCR products were cloned into pBluescript, verified by DNA sequencing to encode the correct product and then inserted into the *Not*I site at the C-terminus of luciferase⊿SKL in pLUH206 to produce in frame gene fusions. To produce the luciferase⊿SKL-gPGK cXS fusion construct, the *Xho*I to *Sst*I fragment encoding the C-terminus of gPGK was replaced by the homologous fragment encoding the C-terminus of cPGK. The resulting construct encodes a protein with the last 16 amino acids of cPGK replacing the last 36 amino acids of gPGK. The C-terminal serine-serine-leucine of the luciferase–gPGK fusion protein was changed to proline-threonine by a partial digest of the plasmid containing the fusion gene with *Sst*I, blunt-ended with T4 DNA polymerase, followed by digestion with *Stu*I and ligation. The GUS gene was obtained by PCR from the template plasmid pGUSN358 → S (Clontech Laboratories, Inc., Palo Alto, CA) using the primer 5′GCAAAGCT-TATGGGTAACCGTCCTGTAGAAAACCCAACC and the degenerate primer 5′GCGGATCCTTAT(TG)AAAG(TA)(TC)TT(TG)A-TTGTTTGC-CTCCCTGCTGCGGTTT. Of the resulting PCR products, only the sequences encoding wild type GUS and GUS ~ SSL were used in this study. The two fragments were digested with *Hin*dIII and *Bam*HI and inserted into the multicloning site of pTSA-HYG2.

Purified plasmid DNA for the transformation experiments was generated from alkaline lysates as described [10]. Generally, 50–100 μg plasmid DNA was linearized in the tubulin intergenic region by digestion with *Mlu*I and electroporated into procyclic TREU667 cells [10]. Plasmids encoding the GUS gene were linearized at the *Bss*HII site in the tubulin intergenic region. When stable transformants were to be selected, hygromycin B (Calbiochem, La Jolla, CA) was added to 50 μg/ml after three days of growth in Cunningham medium [15]. Digitonin solubilizations were done at increasing concentrations of digitonin as described [10], with 0.2 mg digitonin per mg total cell protein generally resulting in optimal solubilization of cytosolic versus glycosomal proteins, as assayed by marker enzymes [10]. Luciferase activity was assayed in a liquid scintillation counter [10] and GUS activity was determined by incubation of cell extracts with the fluorogenic substrate 4-methyl umbelliferyl glucuronide as described by Jefferson [16].

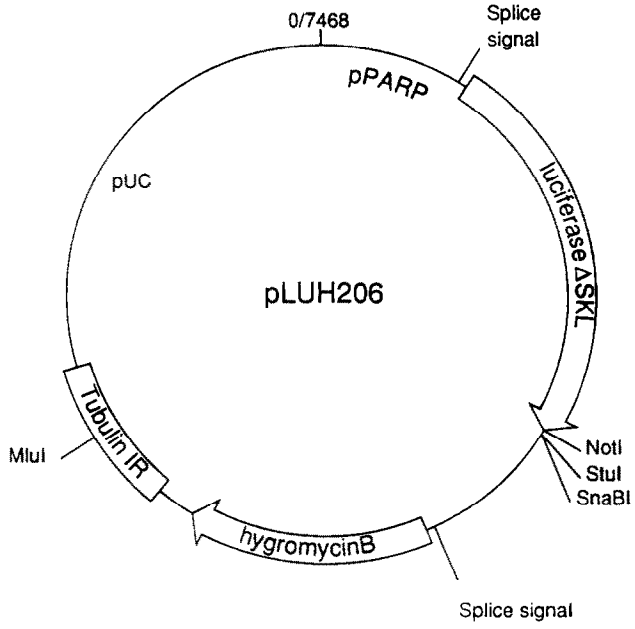

Fig. 1. The transformation vector used for production of C-terminal luciferase fusion proteins. The vector pLUH206 is transcribed by the *T. brucei* promoter for the procyclic acidic repetitive protein (PARP). Following a splice acceptor signal, which is required for *trans*-splicing of messenger RNA in trypanosomes [31,32], we inserted the firefly luciferase gene form which the C-terminal targeting signal (SKL) has been deleted. Unique *Not*I, *Stu*I and *Sna*BI sites at the end of the luciferase gene were used for inserting DNA fragments to construct in frame fusion genes. Preceded by a second splice signal, the hygromycin B gene is used as a selectable marker for the production of stable transformants following homologous integration of the linearized vector into the tubulin locus of *T. brucei*.

Thin sections of stable transformants expressing GUS fusion proteins were prepared according to Tokuyasu [17] and immunolabelled with an anti-β-glucuronidase antibody as described by Keller et al. [18,19]. The antibody was purified as an IgG fraction from a rabbit anti-β-glucuronidase antiserum (Gibco BRL, Gaithersburg, MD) by affinity chromatography on a protein A-Sepharose column [20]. The immunolabelled sections were observed without poststaining in a Philips CM12 transmission electron microscope equipped with a 10 μm diameter objective aperture.

3. RESULTS

3.1. *The C-terminal SSL tripeptide is essential for targeting of gPGK to the glycosomes*

The observation that firefly luciferase is localized to the glycosomes of *T. brucei* and the subsequent identification of its C-terminal tripeptide serine-lysine-leucine (SKL) as a glycosomal targeting signal [10] suggested that this enzyme protein could be used as a marker for glycosomal protein import studies in *T. brucei*. In order to identify the signal(s) responsible for targeting gPGK to the glycosomes, we replaced the C-terminal SKL sequence of luciferase with selected gPGK sequences, thus producing constructs which encode fusion proteins of luciferase and various gPGK sequences. To deter-

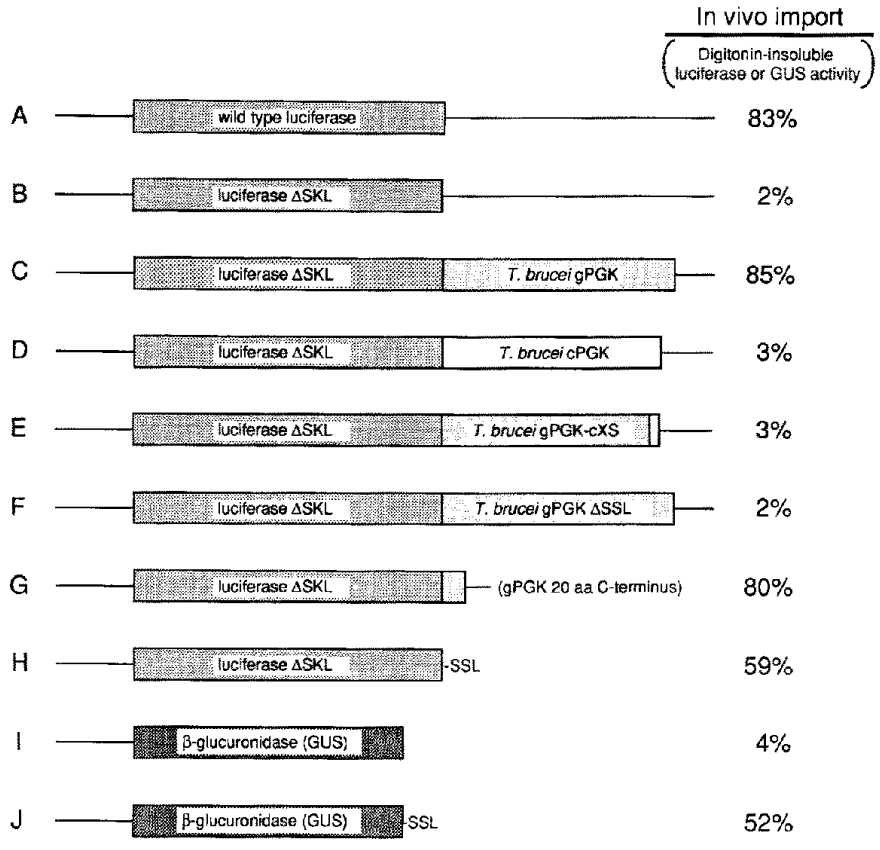

Fig. 2. Luciferase and GUS protein fusions. The fusion constructs shown here are derivatives of the pLUH206 transformation vector and were expressed in procyclic TREU667 cells. Luciferase or GUS assays were were performed 20–72 h after electroporation. The average percentage of enzyme activity associated with the pellet fraction after digitonin extraction and centrifugation is indicated in each case: (A) wild-type luciferase; (B) the C-terminal tripeptide (SKL) was deleted from luciferase; (C) the luciferaseΔSKL gene was fused in frame with the coding region of glycosomal phosphoglycerate kinase (gPGK); (D) the cytoplasmic phosphoglycerate kinase (cPGK) was fused to luciferase; (E) the last 36 amino acids of gPGK were replaced by the last 16 amino acids of cPGK; (F) the serine-serine-leucine (SSL) sequence at the C-terminus of gPGK was replaced by proline-threonine; (G) the 20 amino acid C-terminal extension of gPGK was fused to the C-terminus of luciferaseΔSKL; (H) the SKL sequence at the C-terminus of luciferase was replaced with SSL; (I) wild-type β-glucuronidase; (J) the C-terminal tripeptide of gPGK (SSL) was fused to the C-terminus of β-glucuronidase.

mine the import efficiency of such fusion proteins, we selectively permeabilized the plasma membrane with low concentrations of digitonin to release cytoplasmic proteins. Under the conditions used, the glycosomal membranes remained intact and the glycosomes could be pelleted by centrifugation. We have shown previously that in transformed cells, the ratio of digitonin-solubilized to total luciferase activity agrees well with the proportion of luciferase imported into the glycosome determined by immunoelectron microscopy [10].

In order to simplify the construction of fusion genes and deletion analysis, the C-terminus of the luciferase gene was altered to delete the SKL targeting signal (luciferaseΔSKL) and to include the unique restriction sites *Not*I, *Stu*I and *Sna*BI, as well as stop codons in each reading frame in our expression vector pLUH206 (Fig. 1). Inclusion of the *Not*I site also required the change of the lysine247 immediately preceding the SKL sequence to an arginine. This amino acid substitution in the wild type luciferase had no effect on enzymatic activity or localization of the luciferase to the glycosomes of *T. brucei* and should thus not affect the targeting results of our fusion constructs (data not shown). We inserted the full-length gene coding for either gPGK or cPGK into the *Not*I site of this vector and transformed procyclic TREU667 cells by electroporation [10]. Luciferase assays of these transformants indicated that the fusion proteins remained enzymatically active. At least 80% of the luciferase–gPGK fusion protein was recovered in the pellet fraction following digitonin extraction, indicating efficient import of this protein into the glycosome (Fig. 2C). The almost complete extraction of the equivalent cPGK fusion protein into the supernatant fraction indicated that it remained entirely in the cytosol (Fig. 2D).

We took advantage of a unique *Xho*I site that is conserved in both the gPGK and cPGK genes to construct

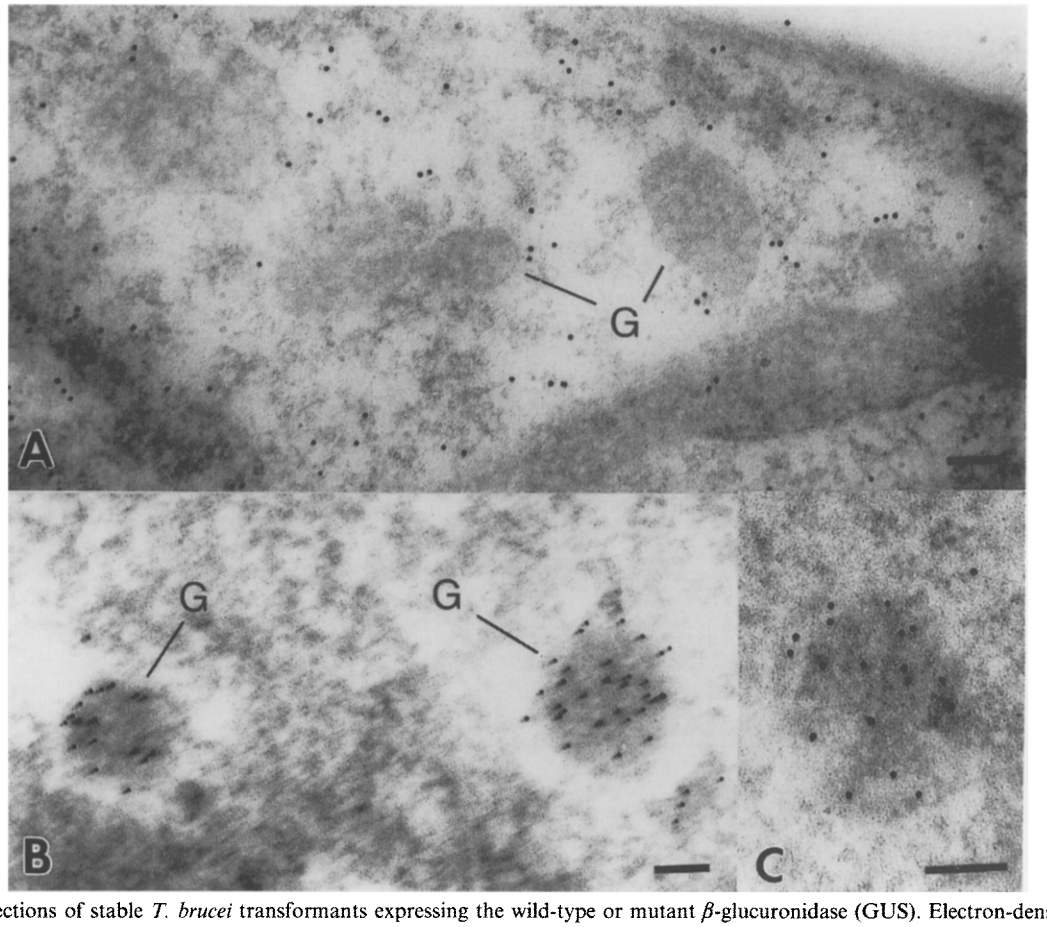

Fig. 3. Cryosections of stable *T. brucei* transformants expressing the wild-type or mutant β-glucuronidase (GUS). Electron-dense gold particles represent the distribution of the GUS protein. (A) Wild-type GUS is primarily localized in the cytoplasm and does not enter the glycosomes (G). (B and C) The GUS ~ SSL fusion protein is imported into the glycosomes. Bars=0.1 μm.

a chimeric PGK gene in which the 3′ end of gPGK is replaced by cPGK sequence. In the resulting fusion protein, the last 36 amino acids of gPGK were replaced with the C-terminal 16 amino acids of cPGK. Since the first 15 amino acids in both of these sequences are conserved, this replacement essentially produced a fusion protein lacking the 20 amino acid C-terminal extension of gPGK. This fusion protein (gPGK-cXS) expressed in the transformed cells was completely solubilized by digitonin (Fig. 2E). An *Sst*I site located at the C-terminus of gPGK allowed us to change the serine-serine-leucine sequence to proline and threonine. The resulting luciferaseΔSKL–gPGKΔSSL fusion protein was also located in the cytosol (Fig. 2F). Thus, the SSL tripeptide at the C-terminus of gPGK is essential for the import of the fusion protein into the glycosome. These results also argue against the presence of a second, independently competent glycosomal targeting signal within gPGK.

To determine whether the C-terminal extension of gPGK alone contains sufficient information for glycosomal targeting we added the last 20 amino acids of gPGK to the C-terminus of luciferaseΔSKL in the pLUH206 construct. At least 80% of the resulting fusion protein was transported to the glycosome (Fig. 2G). By anaolgy to the minimal peroxisomal import signal of luciferase (SKL), we added just the SSL tripeptide to the C-terminus of luciferaseΔSKL (Fig. 2H). 59% of the total luciferase activity was recovered in the pellet fraction following digitonin extraction, indicating that the import efficiency of this protein was still about 70% of those of the wild-type luciferase and gPGK fusion constructs.

3.2. *SSL is sufficient for targeting β-glucuronidase to the glycosome*

There has been some evidence suggesting that sequences in the N-terminal portion of firefly luciferase may also be required for import into the peroxisomes of *Saccharomyces cerevisiae* [21]. If targeting information in addition to the C-terminal tripeptide is also required for import of proteins into the glycosome, the luciferase sequence must contain such an internal targeting signal which can substitute for a similar signal also present in gPGK. We thus used a cytoplasmic pro-

```
Luciferase~SSL  - D A R K I R E I L I K A K K G G K S S L *
                                                : : :
gPGK            - A V V S Y A S A G T G T L S N R W S S L *
                                                . : : :
GUS~SSL         - R W T G M N F G E K P Q Q G G K Q S S L *
```

Fig. 4. C-Termini of the firefly luciferase ~ SSL, *T. brucei*, gPGK and *E. coli* β-glucuronidase~SSL proteins. Structural features of these sequences (see Discussion) were analyzed using the MacVector sequence analysis software (International Biotechnologies, Inc., New Haven, CT).

tein, *E. coli* β-glucuronidase (GUS), to determine if a C-terminal SSL sequence is sufficient for targeting and import of this protein into the glycosome. It is known, that both C-terminal and N-terminal protein fusions can be made with GUS that have little effect on GUS activity [22]. With proper targeting signals, such fusion proteins can be translocated across nuclear membranes [23] or imported into endoplasmic reticulum [24], mitochondria [25] or chloroplasts [26].

Digitonin-solubilization experiments of transformants expressing the GUS protein revealed that it is localized exclusively in the cytoplasm. In those transformants producing the mutant GUS protein with an added serine-serine-leucine tail (GUS ~ SSL), at least 50% of the total GUS activity entered into the pellet fraction (Fig. 2J), which is similar to the results obtained with the luciferase ~ SSL construct mentioned above. The import of GUS ~ SSL into the glycosome was verified by immunolocalization of the protein in frozen thin sections of stable transformants, followed by electron microscopic analysis. The results showed that while all of the wild-type GUS protein remained in the cytoplasm (Fig. 3A), most of the GUS ~ SSL protein was inside the glycosomes of the transformed cells (Fig. 3B and C).

4. DISCUSSION

The signals involved in targeting of proteins to the glycosome have been a matter of considerable speculation in the past [13,27,28]. The finding that most of *T. brucei* glycosomal enzymes contain a high number of basic amino acid residues when compared to their cytosolic or mammalian counterparts led to the proposal that these additional positive charges, which consist of two clusters ('hotspots') separated by 40 Å; on the surface of the presumed three-dimensional protein structures, may be involved in the import of these enzymes into the glycosome [13]. This hypothesis is supported by the observation that a third *T. brucei* PGK (PKG-A), which is also targeted to the glycosome [29,30], is highly homologous to the gPGK protein, but lacks the C-terminal extension found in gPGK. One of the two positively charged hotspots found in gPGK is conserved in PGK-A and, in addition, this protein contains a unique 80 amino acid insertion rich in basic amino acid residues. The *T. brucei* gPGK and PGK-A proteins thus have a predicted isoelectric point (p*I*) of 9.3 and 9.0, respectively, as opposed to 6.8 for cPGK. These charge differences could be important for targeting of gPGK to the microbody membrane, if not for a direct interaction with a putative receptor protein. However, our results show that the C-terminal tripeptide of gPGK is sufficient to target luciferase, which has an estimated p*I* of 6.4, to the glycosomes in vivo. Moreover, the GUS protein, with a predicted p*I* of 5.2, is also efficiently imported into the glycosome when an SSL tripeptide was added to its C-terminus. We conclude that neither a high p*I* nor the hotspots are required for import of gPGK into the glycosome of *T. brucei*, although we cannot completely rule out the remote possibility that a fortuitously similar arrangement of positive charges on the surface of gPGK, luciferase and GUS contributes to their import into the glycosome. The finding that the hotspots are not conserved in the gPGK of another kinetoplastid, *Crithidia fasciculata* [27], further supports our notion that they are not required for import.

Our data indicate that gPGK, when expressed in the form of a fusion protein, contains only one independently functional targeting signal. From the above studies and our previous results [10], we conclude that the minimal requirement for targeting gPGK to the glycosome is the C-terminal SSL sequence. Additional information within the 20 amino acid extension of gPGK may raise the import efficiency of luciferase ~ SSL, as seen by the increase of pelleted luciferase activity from 59% to 83% or approximately that observed for wild-type luciferase (Fig. 2).

A comparison of the C-termini of the three constructs ending in SSL which were used in this study (Fig. 4) also did not reveal any significant sequence similarity apart from the C-terminal tripeptide that may contribute to glycosomal import. Overall, the gPGK C-terminus is considerably more hydrophobic than the two other sequences and it contains only one charged amino acid, although all three sequences display a net positive charge. Of the three, only the luciferase C-terminus has a small probability of forming a helical structure, with the center portion characteristic of a amphipathic helix. However, there is no evidence from peroxisomal studies, that the secondary structure immediately upstream of of a C-terminal targeting signal has any effect on the import of the protein.

In summary, our results indicate that the only specific sequence requirement for import of gPGK into the glycosomes resides in the C-terminal three amino acids serine-serine-leucine. It is a sufficient requirement for glycosomal import of the three vastly different proteins gPGK, luciferase and GUS. It cannot be ruled out, however, that the region immediately upstream of this C-terminal tripeptide in the 20 amino acid C-terminal extension of gPGK may have a minor function, perhaps

related to the presentation of the targeting signal on the surface of the protein or to enhance the interaction with a putative receptor protein.

Acknowledgements: We thank Li Shen for excellent technical assistance. This work was supported by Public Health Service Grants AI-21786 to C.C.W. and AI-22635 to M.P.

REFERENCES

[1] Opperdoes, F.R. and Borst, P. (1977) FEBS Lett. 80, 360–364.
[2] Fairlamb, A.H. and Opperdoes, F.R. (1986) in: Carbohydrate Metabolism in Cultured Cells (Morgan, M.J. ed.) pp. 183–224, Plenum Press, New York/London.
[3] Parsons, M. and Nielsen, B. (1990) Exp. Parasitol. 70, 276–285.
[4] Opperdoes, F.R. (1987) Annu. Rev. Microbiol. 41, 128–151.
[5] Opperdoes, F.R., Baudhuin, P., Coppens, I., De Roe, C., Edwards, S.W., Weijers, P.J. and Misset, O. (1984) J. Cell Biol. 98, 1178–1184.
[6] Borst, P. (1986) Biochem. Biophys. Acta 866, 179–203.
[7] Hart, D., Baudhuin, P., Opperdoes, F.R. and deDuve, C. (1987) EMBO J. 6, 1403–1411.
[8] Gould, S.J., Keller, G.-A., Hosken, N., Wilkinson, J. and Subramani, S. (1989) J. Cell Biol. 108, 1657–1664.
[9] Fung, K. and Clayton, C. (1991) Mol. Biochem. Parasitol. 45, 261–264.
[10] Sommer, J.M., Cheng, Q.-L., Keller, G.-A. and Wang, C.C. (1992) Mol. Biol. Cell 37, 49–759.
[11] Keller, G.-A., Gould, S.J., DeLuca, M. and Subramani, S. (1987) Proc. Natl. Acad. Sci. USA 84, 3264–3268.
[12] Osinga, K.A., Swinkels, B.W., Gibson, W.C., Borst, P., Veeneman, G.H., Van Boom, J.H., Michels, P.A.M. and Opperdoes, F.R. (1985) EMBO J. 4, 3811–3817.
[13] Wierenga, R.K., Swinkels, B., Michels, P.A.M., Osinga, K., Misset, O., Van Beeumen, J., Gibson, W.C., Postma, J.P.M., Borst, P. and Opperdoes, F.R. (1987) EMBO J. 6, 215–221.
[14] Alexander, K., Parail, A.C. and Parsons, M. (1990) Mol. Biochem. Parasitol. 42, 293–296.
[15] Cunningham, I. (1977) J. Protozool. 24, 325–329.
[16] Jefferson, R.A. (1987) Plant Mol. Biol. Rep. 5, 387–405.
[17] Tokuyasu, K.T. (1973) J. Cell Biol. 57, 551–565.
[18] Keller, G.-A., Tokuyasu, K.T., Dutton, A.H. and Singer, S.J. (1984) Proc. Natl. Acad. Sci. USA 81, 5744–5747.
[19] Keller, G.-A., Krisans, S., Gould, S.J., Sommer, J.M., Wang, C.C., Schliebs, W., Kunau, W., Brody, S. and Subramani, S. (1991) J. Cell Biol. 114, 893–904.
[20] Harlow, E. and Lane, D. (1988) Antibodies - a Laboratory Mmanual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[21] Distel, B., Gould, S.J., Voorn-Brouwer, T., van der Berg, M., Tabak, H.F. and Subramani, S. (1992) The New Biol. 4, 157–165.
[22] Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) EMBO J. 6, 3901–3907.
[23] Jarvis, D.L., Bohlmeyer, D.A. and Garcia Jr., A. (1991) Virology 185, 795–810.
[24] Pang, S.Z., Rasmussen, J., Ye, G.N. and Sanford, J.C. (1992) Gene 122, 229–234.
[25] Schmitz, U.K., Lonsdale, D.M. and Jefferson, R.A. (1990) Current Genetics 17, 261–264.
[26] Kavanagh, T.A., Jefferson, R.A. and Bevan, M.W. (1988) Mol. Gen. Genet. 215, 38–45.
[27] Swinkels, B.W., Evers, R. and Borst, P. (1988) EMBO J. 7, 1159–1165.
[28] Kendall, G., Wilderspin, A.F., Ashall, F., Miles, M.A. and Kelly, J.M. (1990) EMBO J. 9, 2751–2758.
[30] Swinkels, B.W., Loiseau, A., Opperdoes, F.R. and Borst, P. (1992) Mol. Biochem. Parasitol. 50, 69–78.
[31] Parsons, M., Nelson, R.G., Watkins, K.P. and Agabian, N. (1984) Cell 38, 309–316.
[32] De Lange, T., Michels, P.A.M., Veerman, H.J.G., Cornelissen, A.W.C.A. and Borst, P. (1984) Nucleic Acids Res. 12, 3777–3790.